

To Denoise or Deblur: Parameter Optimization for Imaging Systems

Kaushik Mitra^a, Oliver Cossairt^b and Ashok Veeraraghavan^a

^aElectrical and Computer Engineering, Rice University, Houston, TX 77005

^bElectrical Engineering and Computer Science, Northwestern University, Evanston, IL 60208

ABSTRACT

In recent years smartphone cameras have improved a lot but they still produce very noisy images in low light conditions. This is mainly because of their small sensor size. Image quality can be improved by increasing the aperture size and/or exposure time however this make them susceptible to defocus and/or motion blurs. In this paper, we analyze the trade-off between denoising and deblurring as a function of the illumination level. For this purpose we utilize a recently introduced framework for analysis of computational imaging systems that takes into account the effect of (1) optical multiplexing, (2) noise characteristics of the sensor, and (3) the reconstruction algorithm, which typically uses image priors. Following this framework, we model the image prior using Gaussian Mixture Model (GMM), which allows us to analytically compute the Minimum Mean Squared Error (MMSE). We analyze the specific problem of motion and defocus deblurring, showing how to find the optimal exposure time and aperture setting as a function of illumination level. This framework gives us the machinery to answer an open question in computational imaging: To deblur or denoise?.

1. INTRODUCTION

In recent years smartphone cameras have improved a lot, starting from the meagre 2 mega-pixels camera of the first generation iPhone to the 41 mega-pixels camera of Nokia 808 PureView. The sensor size has also improved from 1/4" of iPhone 3GS to 1/1.2" of Nokia 808 PureView. However, even with these improvements the image quality of the smartphones are much inferior to that of the current DSLR cameras, especially in low light conditions. This is not much surprising given that the full-frame DSLR camera sensor is about 3 times bigger than the 1/1.2" sensor of Nokia PureView. Apart from a bigger pixel size, another way to improve the image quality (reduce noise) in the low light conditions is to increase the aperture and/or exposure time. However, this make the images susceptible to defocus and/or motion blurs. Thus, we have to deal with a trade-off between denoising and deblurring. We can either capture a noisy image and denoise it later or we can capture a blurred (less noisy) image and deblur it later. The trade-off between denoising and deblurring of course depends on the light level. In high light level, we may prefer to capture the image with a small aperture and/or exposure time, so that the image is sharp and we can avoid the deblurring step, whereas in low light level we might prefer to capture the image with large aperture and/or exposure time so that we have a less noisy image. In this paper, we analyze the denoising vs. deblurring trade-off as a function of light level.

For our analysis we use a recently introduced comprehensive framework for analyzing computational imaging (CI) systems,¹ which simultaneously takes into account the effect of multiplexing, sensor noise, and signal priors. The effect of multiplexing and noise have been studied extensively in the literature.²⁻⁵ However, it has historically been very difficult to determine exactly how much of an increase in performance to expect from signal priors, making it difficult to provide a fair comparison between different cameras. In Mitra et al.,¹ the signal prior is modeled using a Gaussian Mixture Model (GMM) which has two unique properties: : Firstly, GMM satisfies the universal approximation property which says that any probability density function (with a finite number of discontinuities) can be approximated to any fidelity using a GMM with an appropriate number of mixtures.⁶ Secondly, the GMM prior is analytical tractable which allows us to derive simple expressions for the MMSE. We use the analytic expression of MMSE derived in⁷ to study the denoising vs deblurring trade-off as a function of light level.

In this work, we follow a line of research whose goal is to relate performance of the imaging systems to practical considerations (e.g. illumination conditions and sensor characteristics). Following the convention adopted in,^{4,8,9} we define a conventional camera as an impulse imaging system, which measures the desired signal directly (e.g. without blur). Defocus and motion deblurring systems performance are then compared against the corresponding impulse imaging system. Noise is related to the lighting level, scene properties and sensor characteristics. Defocus and motion deblurring

E-mails: Kaushik.Mitra@rice.edu, ollie@eecs.northwestern.edu and vashok@rice.edu

cameras capture blurry images, and all-focused images are then recovered computationally via deconvolution. We consider a pillbox shaped blur function for defocus blur, and a 1-D rect function for motion blur. The impulse imaging counterpart for defocus blur is a narrow aperture image. For motion blur, the impulse imaging counterpart is a short exposure image. Deblurring systems capture more light, but they require deconvolution, which amplifies noise. Impulse imaging doesn't require deconvolution, but captures less light. There is a parameter for both motion (exposure time) and defocus deblurring (aperture size) that can be adjusted to tradeoff light gathering power and deconvolution noise. We address the problem of how to optimize this parameter to achieve the best possible performance with signal priors taken into account.

2. RELATED WORK

Several recent works have analyzed the performance of multiplexed acquisition techniques, including the seminal work by Harwit and Sloan,² as well as more recent techniques that incorporate signal-dependent noise into the analysis.^{3-5,10,11} Recently, Cossairt et. al.⁹ have analyzed CI systems taking into consideration the application (e.g. defocus deblurring or motion deblurring), lighting condition (e.g. moonlit night or sunny day), scene properties (e.g. albedo, object velocity) and sensor characteristics (size of pixels). However, all the above works, do not analyze the performance of CI systems when a signal prior is used for demultiplexing. A few recent papers have analyzed the fundamental limits of the performance of image denoising in the presence of image priors.^{12,13} A similar approach was used by Mitra et al.¹ to extend this analysis to general framework for analyzing computational imaging systems. This framework has been used in this paper for analyzing motion and defocus deblurring cameras. We analyze only single image CI techniques. Multi-image capture techniques have been analyzed by Hasinoff et al.¹⁰ (defocus deblurring), and Zhang et al.¹¹ (motion blurring).

3. IMAGING SYSTEM SPECIFICATION

We use assume a linear imaging model that takes into account signal-dependent noise and models signals using a GMM prior.

3.1 Image Formation Model

We consider linear multiplexed imaging systems that can be represented as

$$y = Hx + n, \tag{1}$$

where $y \in R^N$ is the measurement vector, $x \in R^N$ is the unknown signal we want to capture, H is the $N \times N$ multiplexing matrix and n is the observation noise. For the case of 1D motion blur, the vectors x and y represent a scanline in a sharp and blurred image patch, respectively. The multiplexing matrix H is a Toeplitz matrix where the rows contain the system point spread function. For the case of 2D defocus blur, the vectors x and y represent lexicographically reordered image patches, and the multiplexing matrix H is block Toeplitz.

3.2 Noise Model

To enable tractable analysis, we use an affine noise model.^{10,11,14} We model signal independent noise as a Gaussian random variable with variance σ_r^2 . Signal dependent photon noise is Poisson distributed with parameter λ equal to the average signal intensity at a pixel J . We approximate photon noise by a Gaussian distribution with mean and variance λ . This is a good approximation when λ is greater than 10. We also drop the pixel-wise dependence of photon noise and instead assume that the noise variance at every pixel is equal to the average signal intensity.

3.3 Signal Prior Model

We characterize the performance of CI systems under a GMM prior which has two unique properties: Firstly, GMM satisfies the universal approximation property which says that any probability density function can be approximated to any fidelity using a GMM with an appropriate number of mixtures.⁶ Secondly, a GMM prior lends itself to analytical tractability so that we can use MMSE as a metric to characterize the performance of both impulse and CI systems.

3.4 Performance Characterization

We characterize the performance of multiplexed imaging systems under (a) the noise model described in section 3.2 and (b) the scene prior model described in section 3.3. For a given multiplexing matrix H , we will study two metrics of interest: (1) $mmse(H)$ (see section 4 of¹) and (2) multiplexing SNR gain $G(H)$ defined as the SNR gain (in dB) of the multiplexed system H over that of the impulse imaging system I :

$$G(H) = 10\log_{10}\left(\frac{mmse(I)}{mmse(H)}\right). \quad (2)$$

4. COMMON FRAMEWORK FOR ANALYSIS

We study the performance of defocus and motion deblurring systems under the practical consideration of illumination conditions and sensor characteristics. Defocus and deblurring systems improve upon the impulse imaging systems (which directly captures the sharp but noisy image) by allowing more light to be captured by the sensor. However, captured images then require deblurring, which typically results in noise amplification. To improve upon performance, the benefit of increased light throughput needs to outweigh the degradations caused by deblurring. The combined effect of these two processes is measured as the SNR gain. To calculate the SNR gain, we need to first define an appropriate baseline. Following the approach of,⁹ we choose impulse imaging as the baseline for comparison. This corresponds to a traditional camera with a stopped down aperture (for defocus deblur systems) or a short exposure duration (for motion deblur system)

4.1 Scene Illumination Level and Sensor Characteristics

The primary variable that controls the SNR of impulse imaging is the scene illumination level. As discussed in section 3.2, we consider two noise types: photon noise (signal dependent) and read noise (signal independent). Photon noise is given by the average signal level J (which is directly proportional to the scene illumination level), whereas, read noise is independent of it. At low illumination levels read noise dominates the photon noise, whereas, at high illumination levels photon noise dominates the read noise. We compare the (defocus and motion) deblurring systems to impulse imaging systems over a wide range of scene illumination levels.

Given the scene illumination level I_{src} (in lux), the average scene reflectivity (R) and the camera parameters such as the f-number ($F/\#$), exposure time (t), sensor quantum efficiency (q), and pixel size (δ), the average signal level in photo-electrons (J) of the impulse camera is given by^{9*}:

$$J = 10^{15}(F/\#)^{-2}tI_{src}Rq(\delta)^2. \quad (3)$$

In our experiments, we assume an average scene reflectivity of $R = 0.5$ and sensor quantum efficiency of $q = 0.5$, pixel size $\delta = 5\mu m$ and we set the aperture at $F/11$ and exposure time at $t = 6$ milliseconds. We also assume a sensor read noise of $\sigma_r = 4e^-$ which is typical for today’s CMOS sensors.

4.2 Experimental Details

The details of the experimental setup are as follows

- **Learning:** We learn GMM patch priors from a large collection of about 50 million training patches. For learning we use a variant of the Expectation Maximization approach to ascertain the model parameters. For defocus deblurring, we learn a GMM patch prior, of patch size 16×16 , with 1770 Gaussian mixtures. For motion deblurring, we learn a GMM patch prior of patch size 4256 , with 1900 Gaussian mixtures.
- **Analytic Performance metric:** We compare the (defocus and motion) deblurring systems and impulse imaging systems under signal prior. For that we use the analytic lower bound for MMSE (Eqn. (17) in¹). Once the MMSE is computed for the impulse and deblurring systems, we compute the multiplexing SNR gain in dB using Eqn. (2).

*The signal level will be larger for defocus and motion deblurring systems. The increase in signal is encoded in the multiplexing matrix H .

- Simulations Results for Comparison:** In order to validate our analytic predictions, we also performed extensive simulations. In our simulations, we used the MMSE estimator (Eqn. (12) in¹) to reconstruct the original (sharp) image. The MMSE estimator has been shown to provide state of art results for image denoising,¹⁵ and here we extend these powerful methods for general demultiplexing.

5. OPTIMAL PARAMETER SETTING FOR MOTION AND DEFOCUS DEBLURRING

Optimal exposure setting for motion deblurring

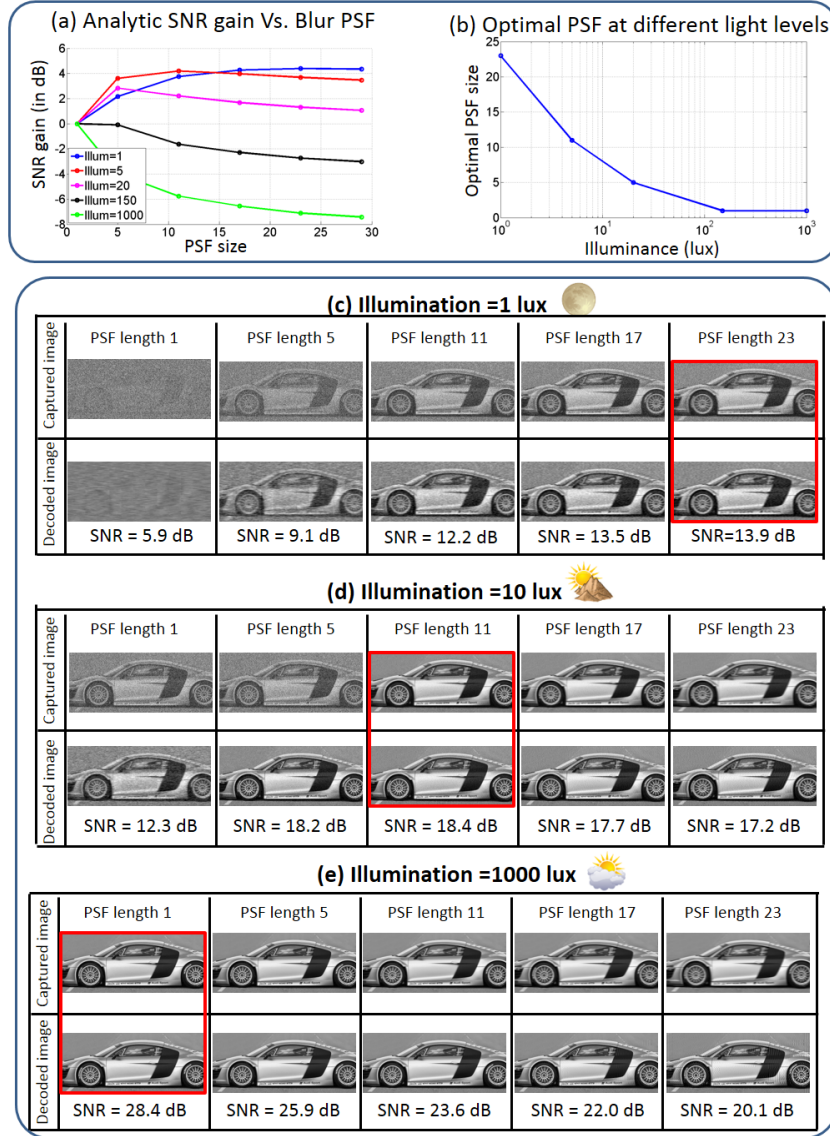


Figure 1. *Optimal exposure setting for motion deblurring:* In subplot (a) we analytically compute the expected SNR gain of different exposure settings (PSF kernel lengths) with respect to the impulse imaging system (of PSF kernel length 1) at various light levels. The impulse imaging system has an exposure time of 6 milliseconds. Subplot (b) shows the optimal blur PSF length at different light levels. At light level of 1 lux, the optimal PSF size is 23 (corresponding to an exposure time of 138 milliseconds), whereas for light levels greater than or equal to 150 lux the optimal is size 1, i.e., the impulse image setting. Subplots (c-e) show the simulated results with different PSF kernel lengths at a few lighting levels.

We use the analytic lower bound for MMSE (Eqn. (17) in¹) to compute the optimal parameter setting of conventional

camera for motion and defocus deblurring. The optimal parameters, exposure for motion deblurring and aperture for defocus deblurring, obviously depends on the scene light level. Hence, we compute the optimal parameters as a function of light levels.

5.1 Optimal Exposure Setting for Motion Deblurring

We first fix the exposure setting of the impulse imaging system based on the range of object velocities we want to capture. The exposure is set to a value that the motion blur for the desired range of velocities is less than a pixel. We then analytically compute the expected SNR gain of different exposure settings (PSF kernel lengths) with respect to the impulse imaging system (of PSF kernel length 1) at various light levels, see Figure 1(a). For light levels less than 150 lux capturing the image with a larger exposure and then deblurring is a better option, whereas, for light levels greater than 150 lux we should capture the impulse image and then denoise. The impulse imaging system has an exposure time of 6 milliseconds. Figure 1(b) shows the optimal blur PSF length at different light levels. At light level of 1 lux, the optimal PSF length is 23 (corresponding to an exposure time of 138 milliseconds), whereas for light levels greater than or equal to 150 lux the optimal length is 1, i.e., the impulse image setting. Figure 1 (c-e) show the simulated results with different PSF kernel lengths at a few lighting levels.

5.2 Optimal Aperture Setting For Defocus Deblurring

Depending on the desired DOF, we fix the aperture size of the impulse imaging system so that the defocus blur is less than a pixel. We then analytically compute the SNR gain of different aperture settings (PSF kernel size) with respect to impulse imaging system of PSF kernel size 1×1 (corresponding to an aperture setting of $F/11$) for various light levels, see Figure 2(a). From this plot, we conclude that for light levels less than 400 lux capturing the image with a larger aperture and then deblurring is a better option, whereas, for light levels greater than 400 lux we should capture the impulse image and then denoise. Figure 2(b) shows the optimal blur PSF size at different light levels. At light level of 1 lux, the optimal PSF is 9×9 (corresponding to an aperture setting of $F/1.2$), whereas, for light levels greater than 400 lux the optimal is 1×1 , i.e., the impulse image setting. Figure 2(c-d) show the simulated results with different PSF size at a few lighting levels.

6. DISCUSSIONS

We have analysed the problem of parameter optimization for motion and defocus deblurring cameras, answering the question: To deblur or denoise?. We use a GMM to model the signal prior and compute performance for deblurring and impulse imaging systems. We relate performance to illumination conditions and sensor characteristics. We showed for a typical camera specification, denoising is preferable when the illumination is greater than 150 lux for motion deblurring, and 400 lux for defocus deblurring. In addition, we optimized parameters (exposure time for motion blur, aperture size for defocus blur) at dimmer illumination conditions.

REFERENCES

- [1] Mitra, K., Cossairt, O., and Veeraraghavan, A., "A framework for the analysis of computational imaging systems with practical applications," *CoRR* **abs/1308.1981**, **abs/1308.1981** (2013). 1, 2, 3, 4
- [2] Harwit, M. and Sloane, N., "Hadamard transform optics," *New York: Academic Press* (1979). 1, 2
- [3] Wuttig, A., "Optimal transformations for optical multiplex measurements in the presence of photon noise," *Applied Optics* **44** (2005). 1, 2
- [4] Ratner, N. and Schechner, Y., "Illumination multiplexing within fundamental limits," in [CVPR], (2007). 1, 2
- [5] Ratner, N., Schechner, Y., and Goldberg, F., "Optimal multiplexed sensing: bounds, conditions and a graph theory link," *Optics Express* **15** (2007). 1, 2
- [6] Sorenson, H. W. and Alspach, D. L., "Recursive bayesian estimation using gaussian sums," *Automatica* **7**, 465–479 (1971). 1, 2
- [7] Flam, J., Chatterjee, S., Kansanen, K., and Ekman, T., "On mmse estimation: A linear model under gaussian mixture statistics," *Signal Processing, IEEE Transactions on* **60** (2012). 1
- [8] Ihrke, I., Wetzstein, G., and Heidrich, W., "A theory of plenoptic multiplexing," in [CVPR], (2010). 1
- [9] Cossairt, O., Gupta, M., and Nayar, S. K., "When does computational imaging improve performance?," *IEEE transactions on image processing* **22**(1-2), 447–458 (2013). 1, 2, 3
- [10] Hasinoff, S., Kutulakos, K., Durand, F., and Freeman, W., "Time-constrained photography," in [ICCV], 1–8 (2009). 2
- [11] Zhang, L., Deshpande, A., and Chen, X., "Denoising versus Deblurring: HDR techniques using moving cameras," in [CVPR], (2010). 2
- [12] Chatterjee, P. and Milanfar, P., "Is denoising dead?," *Image Processing, IEEE Transactions on* **19**(4), 895–911 (2010). 2

Optimal aperture setting for defocus deblurring

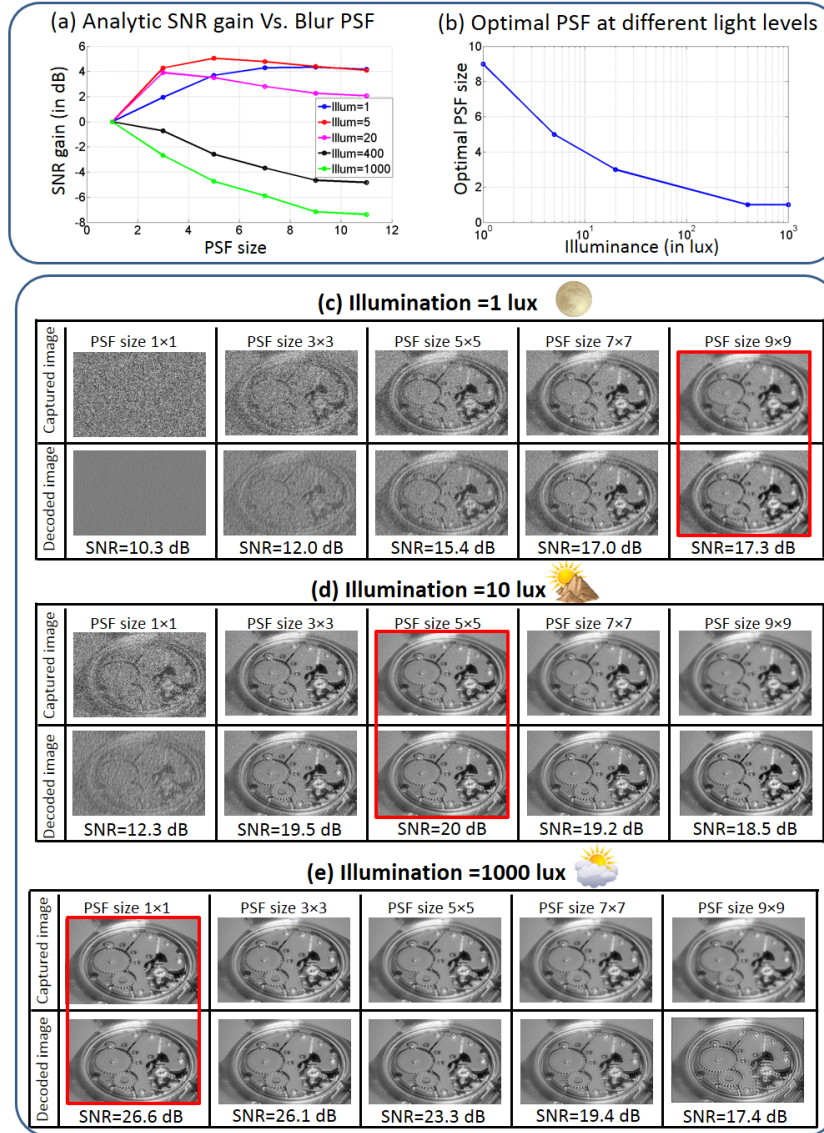


Figure 2. *Optimal aperture setting for defocus deblurring*: In subplot (a) we analytically compute the SNR gain of different aperture settings (PSF kernel size) with respect to impulse imaging system of PSF kernel size 1×1 (corresponding to an aperture setting of $F/11$) for various light levels. In subplot (b) we show the optimal blur PSF size at different light levels. At light level of 1 lux, the optimal PSF is 9×9 (corresponding to an aperture setting of $F/1.2$), whereas for light levels greater than 400 lux the optimal is 1×1 , i.e., the impulse image setting. Subplots (c-d) show the simulated results with different PSF size at a few lighting levels.

- [13] Levin, A., Nadler, B., Durand, F., and Freeman, W. T., “Patch complexity, finite pixel correlations and optimal denoising,” in [ECCV], (2012). 2
- [14] Schechner, Y., Nayar, S., and Belhumeur, P., “Multiplexing for optimal lighting,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **29**(8), 1339–1354 (2007). 2
- [15] Levin, A. and Nadler, B., “Natural image denoising: Optimality and inherent bounds,” in [CVPR], (2011). 4